Adaptabilities of three mainstream short-term wind power forecasting methods

Jie Yan,[1,a)] Xiaoli Gao,[2] Yongqian Liu,[1] Shuang Han,[1] Li Li,[1] Xiaomei Ma,[1] Chenghong Gu,[3] Rohit Bhakar[3] and Furong Li[3]

[1]*State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Changping District, Beijing102206, China*

[2]*Northwest Electric Power Design Institute Co. Ltd. Of China Power Engineering Constitute Group, Xián 710075, China*

[3]*Department of Electronic and Electrical Engineering, University of Bath, Bath BA2 7AY, U.K*

Abstract: Variability and intermittency of wind is the main challenge for making a reliable wind power forecasting (WPF). Meteorological and topological complexity makes it even harder to fit any forecasting algorithm to one particular case. This paper presents comparison of three short term WPF models based on three wind farms in China with different terrains and climates. The sensitivity effects of training samples on forecasting performance are investigated in terms of sample size, sample quality and sample time scale. Then, their adaptabilities and modeling efficiency are also discussed under different seasonal and topographic conditions. Results show that 1) radial basis function (RBF) and support vector machine (SVM) generally have higher prediction accuracy than that of genetic algorithm back propagation (GA-BP), but different models show advantages in different seasons and terrains. 2) WPF taking a month as the training time interval can increase the accuracy of short-term WPF. 3) The change of sample number for GA-BP and RBF is less sensitive than that of SVM. 4) GA-BP forecasting accuracy is equally sensitive to all size of training samples. RBF and SVM have different sensibility to different size of training samples. This study can quantitatively provide reference for choosing the appropriate WPF model and further optimization for specific engineering cases, based on better understanding of algorithm theory and its adaptability. In this way, WPF users can select the suitable algorithm for different terrains and climates to achieve reliable prediction for market clearing, efficient pricing, dispatching etc.

Key words: wind power forecasting, training sample, model efficiency, BP neural network, RBF neural network, support vector machines

I. INTRODUCTION

As the increasing energy adequacy and environmental concerns, large-scale integration of wind power has been contributing more to the world energy supply. After a slowdown in 2013, wind industry set a new record for annual installations in 2014. Globally, 51,477 MW of new wind generating capacity was added in 2014 according to the global wind market statistics released by the Global Wind Energy Council (GWEC). The record-setting figure represents a 44% increase in the annual market, and is a solid sign of the recovery of the industry after a rough patch in the past

[a)] Present address: Department of Electronic and Electrical Engineering, University of Bath, Bath BA2 7AY, UK

few years. Total cumulative installations stand at 369,553 MW at the end of 2014[1].

China is the largest wind power market worldwide and has abundant wind energy resources. Wind power is becoming the third main power supplier in China after thermal power and hydro power. In 2014, Chinese wind power industry maintained a robust growth. 19,810 MW of new wind power installed capacity was added which reaching a new record. Wind power cumulative installation capacity was 96,370 MW which accounting for 7% of the total generation capacity of China and 27% of the global wind power installation. In 2014, wind power had 2.78% share of the total generation in power supply. It is estimated that the wind power capacity will reach 30 GW[2]. However, the intermittent nature of wind generation creates challenges for market trading, load management and the economic operation of power systems[3,4].

Reliable wind power forecasting (WPF) is one of effective ways to mitigate technological and economic impacts on the power system, because it provides direct estimation of wind power availability. Unfortunately, it is difficult to make an accurate prediction, especially in an area with complex topography or meteorological environment. Currently, mean average percentage error of the wind speed forecasting is approximately in the 8-20% range in Chinese wind farms[5,41-45]. Therefore, it is important for wind farm owners, system operators, and energy suppliers to understand how existing WPF models perform in different external conditions and to further improve their prediction accuracy.

A great amount of literature has been devoted to short term WPF using statistical or physical approaches. Statistical models employ weather variables from numerical weather prediction (NWP) model to predict time series of future power generation. The relationships between weather variables and historical power generated at a wind farm are modeled by different artificial intelligent techniques. The algorithms include, for instance, time series analysis[6,7], Kalman filtering[8,9], neural network (NN)[10-19], support vector machine (SVM)[20-22,36],and relevance vector machine (RVM)[23,24], etc. Physical models[25-30] estimate wind condition at reference locations according to the atmospheric behaving law, then transfer to power production through power curve. It is independent of the measured historical data and can be used for both the newly built and the operating wind farms. Commonly used methods are the analytical methods or computational fluid dynamics (CFD) methods. For physical models, analytical method is efficient to compute but difficult to meet accuracy requirements. Conversely, CFD method solves the Navier-Stokes equation in every prediction process and simulates wind flow accurately. As a sacrifice, the computational burden of CFD method makes it challenging to satisfy real-time operation requirements.

Among the statistical methods, artificial NN (ANN) is widely used because of its good prediction generalization, but it is sensitive to the size of training samples and suffers from the over-fitting problem. SVM is an effective machine learning method with small samples and can avoid over-fitting. But it can only use a continuous symmetric kernel with positive integral operator; the number of support vectors grows linearly with the size of training sample sets. RVM

is a nonlinear pattern recognition model with simple structure based on Bayesian theory and marginal likelihood functions. It has no kernel function limitation and is offering excellent generalized prediction performance with even smaller samples.

A number of studies tested and compared the performance of different WPF algorithms[31-35]. The results show that strong improvements of up to about 20% have been obtained by using the data of different NWP models or ensemble models as input data for the WPF models. Artificial neural network and vector machine methods can significantly improve the forecasting accuracy over traditional persistence methods, especially after optimization by genetic algorithm (GA) and particle swarm optimization (PSO). The multi-layer perceptron ANN (MLP ANN) outperforms SVM, regression trees and random forests in short term wind speed and wind power forecasting in Jursa and Rohrig's study[12]. But, SVM for system orders from 1 to 11 had better accuracies than MLP ANN in Mohades's wind speed forecasting results[36]. This above contradiction might come from the fact that each WPF methodology might have different adaptabilities and training characteristics towards various external conditions. So, it requires reproducing the results from existing WPFs in extensive practical examples, particularly in wind farms with complex topography and climate patterns. Moreover, a comprehensive understanding on the forecasting performance in China, the largest and newly born wind power industry center, is needed. And the influence of training samples on forecasting is also needed to investigate.

To mitigate the above problems, this paper investigates the characteristics and adaptabilities of various WPF methodologies in three Chinese wind farms with various meteorological and topological conditions. Three statistical WPF models including back propagation optimized by genetic algorithm (GA-BP), Radial Basis Function (RBF), and SVM are considered for evaluation in three Chinese wind farms with distinctive external conditions. The performances of these models are assessed over three aspects: accuracy, adaptability and computational efficiency. The effects of training samples on forecasting accuracy are also analyzed in terms of their sizes, qualities and time scale. In this way, the most representative training samples can be selected for securing the best prediction performance. This work will help wind power predictors to select the most suitable prediction models for different terrains and climates, and it could also help to improve model performance according to the specific characteristics.

The remainder of this paper is as follows: Section II and section III describe the operational data from wind farms in case study and the NWP tool used to provide meteorological variables for WPF. Section IV and V presents a brief summary of the used WPF methods and their evaluation criteria. Case study in section VI shows the comparison results. Finally, section VII presents a summary of the results.

II. DATA

In order to compare the adaptabilities of three WPF models, three wind farms in China are taken as examples to conduct one day ahead of wind power prediction.

◆No. 1 wind farm (WF1) has 134×1.5 MW wind turbines, and is located in the coast of eastern China with a flat terrain. The data are collected with a 5 min interval, covering 12 months from 2011.5.1 to 2012.4.30.

◆No. 2 wind farm (WF2) has 67×1.5 MW wind turbines, and is with a relatively flat terrain in northeast China. The data are collected with a 5 min interval, covering 12 months from 2011.1.5 to 2011.12.31.

◆No. 3 wind farm (WF3) has 100×1.5 MW wind turbines, and is with a complex terrain in the Chinese northern inland regions. The historical measurements are collected with a 5 min interval, covering 12 months from 2011.1.5 to 2011.12.31. The met mast wind speed data are missing from January to March, due to recording device fault. And no interpolation is conducted because there is no available wind speed measured at other height or other met mast.

The data from wind farms include mean wind speed from met mast, the mean power output of wind farms from SCADA (supervisory control and data acquisition) and NWP of the wind components at the hub height of wind turbines. Met masts in all wind farms are located in the upwind prevailing wind direction, measuring the wind speed and wind direction at hub height and other atmosphere variables (pressure, temperature and humidity) at 10 m. The locations of these met masts were carefully selected and assessed before wind resource assessment and wind turbine locating by wind farm owners. So, the representative and the wake effect of met mast are considered. The historical measurements from all wind farms have a high level of availability over the question period. There are fewer than 4% missing operation data because of the occasional device faults in wind farms. The missing parts are simply deleted and no interpolation process is implemented for the missing data. Among the available data, 80% are considered as candidate training samples and remaining 20% are used as test samples. The historical power data are normalized by the installed capacities of respective wind farms. In this way, the specific characteristics of the wind farms can be masked in order to make a scale-free comparison among various wind farms.

III. NUMERICAL WEATHER PREDICTION

As wind vector is the key factor affecting the power generation characteristics of wind turbines, the WPF models take NWP as input data and power output at wind farms as the training targets. In this paper, the NWP data are produced and maintained at State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), in China. The maintenance of NWP system makes sure it can have a consistent run.

As NWP tool, WRF (weather research and forecasting) model predicts the time series of NWP wind speed and wind direction at met mast location to serve as reference point. It downscales to 6km horizontal resolution. NWP wind speed and wind direction are interpolated from the lowest levels in WRF to a hub height of wind turbines. The time resolution of NWP is the same as SCADA. The WRF model is initialized every day with the prediction of GFS (Global Forecasting System). GFS is released by NCEP (National Centres for Environmental Prediction)

corresponding to the assimilation of atmospheric data at 00:00 GMT (Greenwich Mean Time). This prediction horizon of initial field is 72 hours ahead. The NWP is available at about 7:00 GMT.

The initial field of NWP is from NCEP FNL (Final) Operational Global Analysis data, which are calculated based on 1-degree by 1-degree grids prepared operationally every six hours. They are from the Global Data Assimilation System (GDAS), which continuously collects observational data from the Global Telecommunications System (GTS). The FNLs are made with the same model which NCEP uses in the Global Forecast System (GFS), but the FNLs are prepared about an hour or so after the GFS is initialized. The FNLs are delayed so that more observational data can be used. The GFS is run earlier in support of time critical forecast needs, and uses the FNL from the previous 6 hour cycle as part of its initialization[37].

The analyses are available on the surface, at 26 mandatory (and other pressure) levels from 1000 millibars to 10 millibars, in the surface boundary layer and at some sigma layers, the tropopause and a few others. Parameters include surface pressure, sea level pressure, geopotential height, temperature, sea surface temperature, soil values, ice cover, relative humidity, u- and v-winds, vertical motion, vorticity and ozone[37].

IV. SUMMARY OF WPF METHODS

There are three WPF methods implemented in this paper to evaluate their performance in wind farms mentioned above. The methods are summarized in Table I. For comparability, all methods use the same input variables, training samples and test samples; and the wind power forecasting horizon is 24 hours ahead. The input variables include NWP wind speed, sine of wind direction and cosine of wind direction[38,39,40].

The samples are $\{X_i, P_i\}_{i=1}^{i=N}$, where input variables are $X_i = \{ws_i, wd.\sin_i, wd.\cos_i\}$ and the training target is the power output $\{P_i\}$ in a wind farm. $\{ws_i\}$ is NWP speed; $\{wd.sin_i\}$ is sine of wind direction; $\{wd.cos_i\}$ is cosine of wind direction; N is the sample number.

Since the training objective is to minimize the mean square error (MSE) of the network, the fitness function is defined as the follow[46].

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\frac{(y_i - \hat{y}_i)^2}{2} \tag{1}$$

$$fitness = \frac{1}{1+MSE} \tag{2}$$

where $\hat{y}_i$ is the desired model output or the learning target – the measured wind farm power output. When *fitness* is close to 1, the requirement of network accuracy is thought to be reached.

Table I Summary of WPF methods

Forecasting Algorithm	Implementing wind farm

GA-BP	Back propagation artificial neural network optimized by generic algorithm	WF1~WF3
RBF	Radical basis function artificial neural network	WF1~ WF3
SVM	Support vector machine	WF1~WF3

a. Back propagation artificial neural network optimized by generic algorithm

Back propagation (BP) neural network is a three-layer feed-forward network with a supervised learning mechanism, introduced by Rumelhart, and McCelland in 1986. It is trained based on an error back propagation algorithm and can learn any nonlinear mapping of input to target.

There are two stages in BP learning process, which are input information forward propagation and error backward propagation. During forward stage, model input variables X_i, including NWP wind speed and NWP wind direction, are transferred from the input layer, and then weighted by the hidden layer, finally to the output layer. The output of the BP network is the predicted wind power output calculated by the sigmoid transfer function. This output result is compared with the learning targets, which is the observed wind power. The inevitable training error returns backward along the network path from output layer through hidden layer to input layer so that the network parameters (connection weights and thresholds) of each layer can be adjusted. Network parameters of BP are randomly selected in the initial iteration, and then updated according to error response and improves the forecasting performance of the entire network. However, this random parameter selection might result in inappropriate training scale (excessive training or lack of training), local minimum problem or inefficient convergence, especially with large amount of training samples.

Therefore, GA is used to evaluate the viability of the individual neural nodes and accordingly to direct the BP initial parameter and training process. In this way, the global search capability and the generalization of the BP network can be improved. The modeling process of BP optimization using GA is shown in Fig.1.

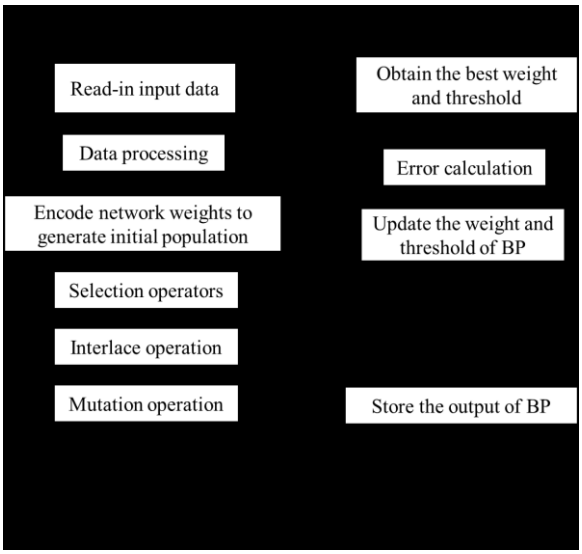

FIG.1. Modeling process of the GA-BP neural network

b. RBF neural network

RBF neural network[47,48] is a three-layer network that learns by measuring the Euclidean distance of data. It consists of an input layer, a single nonlinear hidden layer with nodes, and a linear output layer. There is no weight between input layer and hidden layer. This simple network structure is an advantage of the multilayered perceptron network for efficient training and forecasting.

During the data processing, a hidden node represents one kernel functions and the hidden layer performs nonlinear transforms for the feature extraction. Output layer is a local kernel function with unit normalization which giving a linear and weighted sum of all the hidden nodes outputs. Gaussion function is generally used to quantify the relationship between the model input and the output of the j^{th} node in the hidden layer, as shown in eq.(3). There are two parameters in kernel function determining the range of this kernel function effect, which are center and width. And the expected output of the network is shown in eq.(4).

$$f\left(X_i, C_j, \sigma_j\right) = \exp\left(\frac{-\left\|X_i - C_j\right\|^2}{2\sigma_j^{\,2}}\right) \tag{3}$$

$$y_i = \sum_{j=1}^{M} f\left(X_i, C_j, \sigma_j\right) w_j \tag{4}$$

where $C_j = \left[c_{j1}, c_{j2}, ..., c_{jn}\right]^T$ is the center vector of the j^{th} neuron in the hidden layer; $\sigma = \left[\sigma_1, \sigma_2, ..., \sigma_j\right]^T$ is the width vector; w_j is the weight between the j^{th} neuron of the hidden layer; M is the neuron number in the hidden layer.

The weight parameter updates for minimizing the total fitting error as follow.

$$\min error = \frac{1}{2}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2 \tag{5}$$

In this paper, Matlab toolbox is used to perform the training and testing. The modeling flow

of the RBF model is shown in Fig.2.

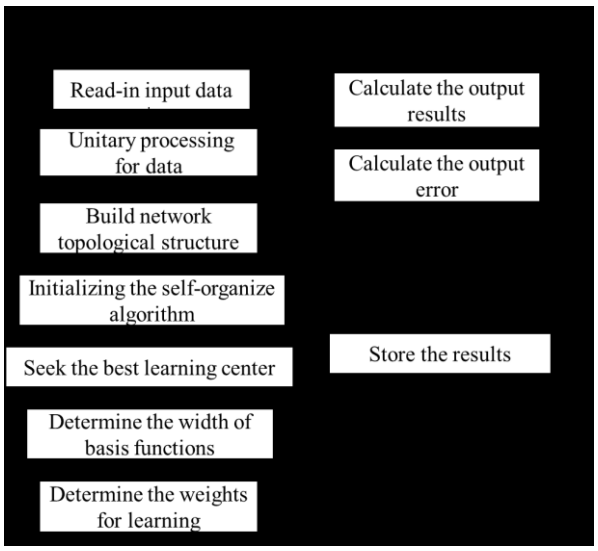

FIG.2. Modeling process of the RBF neural network

C. Support vector machine

The basic idea of support vector machine (SVM) [49] for forecasting is to map the input vector into a high-dimensional feature space by a nonlinear mapping function, and then to perform linear regression in the feature space. The following estimate function is used to make the prediction.

$$f\left(X_i\right)=\left[w\cdot\phi\left(X_i\right)\right]+b \tag{6}$$

where w is weight vector; b is the threshold value; $\phi\left(X_i\right)$ is the nonlinear mapping function.

The function approximating problem is equivalent with the minimization of the risk function R, as shown below.

$$R=\frac{1}{2}\|w\|^2+c\frac{1}{l}\sum_{i=1}^{l}\left|y_i-f\left(X_i\right)\right|_{\varepsilon} \tag{7}$$

where c is the regularized constant determining the trade-off between the empirical error and the regularization term $\frac{1}{2}\|w\|^2$. $|\square|_{\varepsilon}$ is called Vapnik's ε-insensitive loss function and ε is a measure for empirical error[50].

$$\left|y_i-f\left(X_i\right)\right|=\begin{cases}0 & if\ \left|y_i-f\left(X_i\right)\right|\le\varepsilon\\ \left|y_i-f\left(X_i\right)\right|-\varepsilon & otherwise\end{cases} \tag{8}$$

The minimizing function can be written as the following form.

$$f\left(X_i,\alpha_i,\alpha_i^*\right)=\sum_{i=1}^{l}\left(\alpha_i-\alpha_i^*\right)k\left(X_i,X\right)+b \tag{9}$$

where α_i,α_i^* are Lagrange multipliers and $\alpha_i\times\alpha_i^*=0,\ \alpha_i,\alpha_i^*\ge 0$ for any $i=1,2,...,l$. $k\left(x_i,x_j\right)=\phi\left(x_i\right)\times\phi\left(x_i\right)$ is kernel function, which needs to satisfy Mercer's condition. In SVM, it is generally computed by RBF kernel function as follow.

$$k(x,y)=\exp\left(\frac{-\|x-y\|^2}{2\sigma_j^{\,2}}\right) \tag{10}$$

By introducing the slack variables ζ_i,ζ_i^*, the risk function R can be rewritten to the following form. The slack variables are introduced when the data cannot be estimated under the ε.

$$R=\frac{1}{2}\|w\|^2+c\sum_{i=1}^{n}\left(\zeta_i+\zeta_i^*\right) \tag{11}$$

subject to:

$$\begin{cases}\left(w\cdot\phi(x)\right)+b-y_i\le\varepsilon+\zeta_i^* & i=1,2,...,l\\ y_i-\left(w\cdot\phi(x)\right)-b\le\varepsilon+\zeta_i & i=1,2,...,l\\ \zeta_i,\zeta_i^*\ge 0 & i=1,2,...,l\end{cases} \tag{12}$$

In order to obtain the estimates of weights and threshold, a minimization function can be obtained with Lagrange multipliers and Karush-Kuhn-Tucker conditions as follow.

$$\begin{aligned}&Min\quad Lag\left(\alpha^*,\alpha\right)\\&=\varepsilon\sum_{i=1}^{N}\left(\alpha_i^*+\alpha_i\right)-\sum_{i}^{N}y_i\left(\alpha_i^*-\alpha_i\right)+\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\alpha_i^*-\alpha_i\right)\left(\alpha_j^*-\alpha_j\right)k\left(X_i,X_j\right)\end{aligned} \tag{13}$$

subject to:

$$\sum_{i=1}^{N}\left(\alpha_i^*-\alpha_i\right)=0,0\le\alpha_i,\alpha_i^*\le c \tag{14}$$

The data points associating with $\alpha_i,\alpha_i^*\neq 0$ are called support vectors. The key idea for SVM is to select the values of α_i,α_i^* for the data set $\{x_i\}$. That is why the training sample effects on forecasting performance is significant to analyze.

V. EVALUATION CRITERIA

The deterministic WPFs are evaluated and compared on three aspects: prediction accuracy, computation efficiency and model adaptability.

(1) The error criteria adopted in the performance evaluation of the deterministic WPF is Normalized root mean square error (RMSE). It gives a more representative evaluation of the forecasting error for the complete validation period[51,52].

$$RMSE=\frac{\sqrt{\sum_{i=1}^{n}\left(P_{ai}-P_{pi}\right)^2}}{Cap\times\sqrt{n}} \tag{15}$$

where, P_{ai} is actual power output measurement at time of *i*. P_{pi} is predicted power at time of *i*. *Cap* is capacity of wind farm. *n* is number of model training samples.

(2) *Computation efficiency*: The convergence times to train the WPF models are compared. All the cases are implemented on a computer where the configuration is CPU AMD X4830 Processor, 2.80GHz.

(3) *Impacts of training samples*: The impacts of training sample on forecasting accuracy are studied in terms of sample size, sample quality and the time scale of training samples. The training samples are built up with three time scales (year, season, and month).

(4) *Model adaptability*. The adaptabilities of WPF methods in different terrains and climates are also examined by implementing each WPF model in three wind farms with distinct topography and climate patterns.

VI. CASE STUDY

a. Training sample impacts on WPF

The GA-BP, RBF, and SVM models are adopted to conduct the short-term WPF with different training samples in three wind farms. The effects of training samples come from three aspects, as follow: sample number, accuracy of training sample (NWP), and time scale of training samples. The training samples are divided into 36 groups according to NWP speed absolute error and sample number. The summary of training samples is in Fig.3. The nomenclature of sample number for each wind farm is as follow: "Small"-500 samples (green bar); "Medium"-2000 samples (blue bar); "Large"- 5000 samples (purple bar).

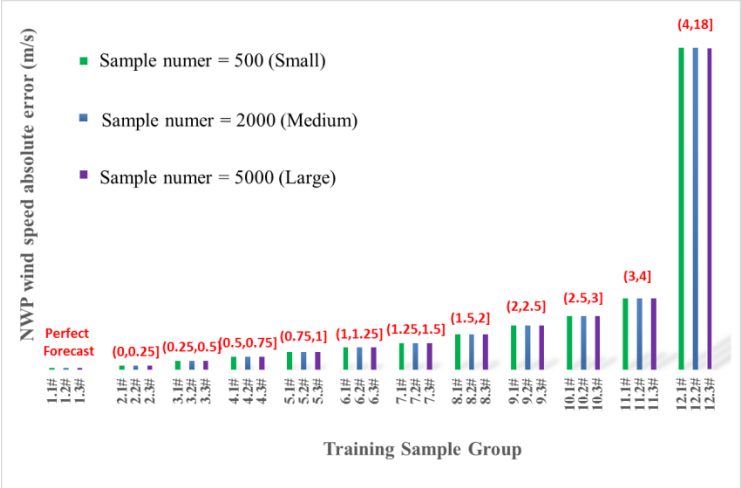

FIG.3 Summary of training samples

Fig.4 shows the forecasting RMSE with different numbers of training samples. To this end, the WPF models are established based on equivalent size of training samples. The blue curve represents samples of 1.1#, 2.1#, 3.1#, 4.1#, …, 12.1#. The red curve represents samples of 1.2#, 2.2#, 3.2#, 4.2#,…, 12.2#. The green curve represents samples of 1.3#, 2.3#, 3.3#, 4.3#,…, 12.3#. X-axis is the representation of NWP wind speed error; and the NWP error is increasing from left to right. The training samples at the same x-axis position have the same NWP wind speed absolute error level. The results show that forecasting with 2000 training samples and 5000 training samples have very similar and relatively stable performance. Blue curves show more variable features, no matter what WPF algorithms are considered. Interesting to see that, the forecasts with perfect NWP (1.1#, 1.2# and 1.3#) do not obtain the best accuracy, although the curves show slightly upward trend with the increase of NWP error level. This verifies the contradiction effects between sample number and NWP accuracy. If the sample number cannot reach the minimum

amount that model training requires, the forecasting accuracy is decreased even the wind speed forecasts are 100% accurate.

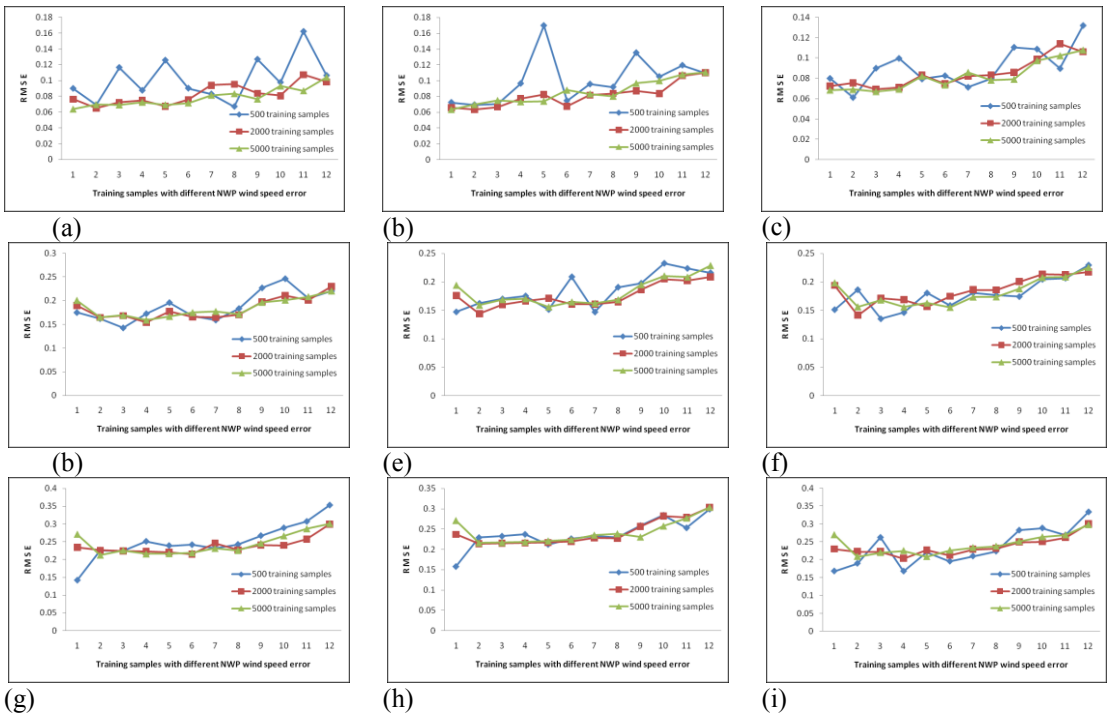

FIG.4 Influence of training sample size on forecasting accuracy. (a) GA-BP in 1#wind farm, (b) RBF in1 #wind farm, (c) SVM in1 #wind farm, (d) GA-BP in2 #wind farm, (e) RBF in2 #wind farm, (f) SVM in2 #wind farm, (g) GA-BP in3 #wind farm, (h) RBF in 3#wind farm, and (i) SVM in3 #wind farm

To manifest the joint influence of sample number and sample accuracy, the forecasting model is established with mixes of NWP with different wind speed forecast error. Fig.5 is the average RMSE value of all wind farm power output forecasting. Specifically, the prediction accuracy of GA-BP increases as the training sample number increases. It means that GA-BP algorithm has larger demand for training samples at this stage. So, it is unsuitable for the modeling of new-built wind farms or areas with a wide variety of weather patterns. But, continuous increase of training samples would inversely lead to forecasting accuracy drops, and then the error levels off. This is because excessive samples arise the over-fitting problem. For RBF, the prediction accuracy also increases with the increase in training sample number. When the number of training samples is less than 500, the rising slope of RBF red curves are far greater than that of GA-BP. But with more than 500 samples, the RBF accuracy remains fairly constant. The prediction accuracy of SVM increases with the decreasing of training samples, however, it reduces rapidly when the training samples are less than 200. This verifies the forecasting capabilities of SVM with small amount of samples.

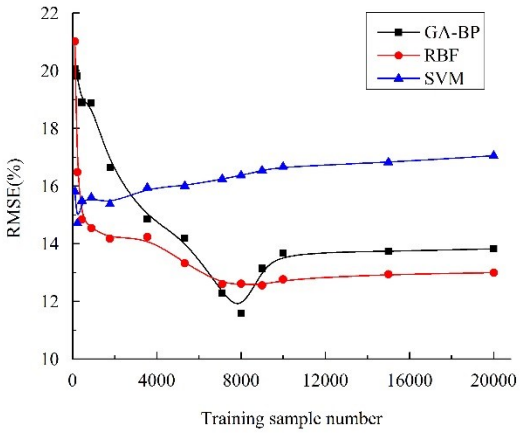

FIG. 5 RMSE of various training sample numbers

In order to investigate the relation between the model accuracy and time scale of training samples, the training data is classified by year, season and month as time intervals. The results of GA-BP, RBF and SVM with different time-scales training and testing are presented in Fig. 6. Constant sample size is used in the plots. The RMSE of the three models increases with the expanding time scale of training samples. It demonstrates that prediction models which are established with month as time interval can improve the prediction accuracy. This is mainly because the changing of the wind might follow specific rules in a certain month or season, which is beneficial for the prediction to build the model for each month or season. In other words, it is hard to grasp the meteorological features from representative samples in a long period of time because the wind and power generation performance would behave in various ways.

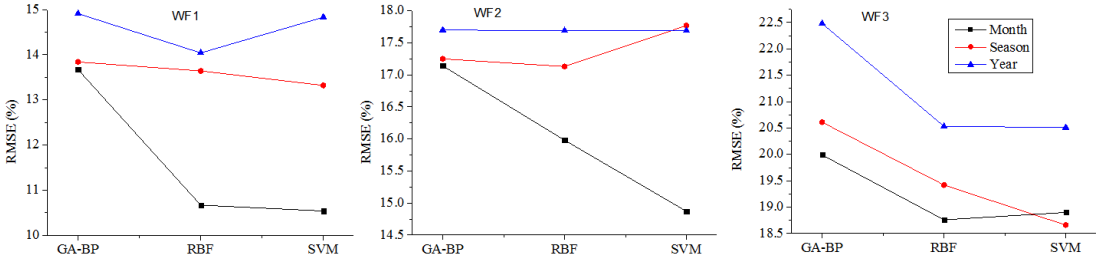

FIG.6 Prediction RMSE with different time scale of training samples

b. Modeling efficiency

Fig.7 is the training time with GA-BP, RBF and SVM based WPF models. Overall, the training time of all WPF algorithms increase as the sample number increases. For GA-BP, the training time gains linearly as the forecasting accuracy linearly increases as well. It indicates that GA-BP sacrifices the modeling efficiency to forecasting accuracy. For RBF, the training time increases exponentially with increase of sample number. It means that RBF is not suitable for training with great amount of samples. Otherwise, its computational efficiency dramatically drops and the forecasting accuracy increases mildly. The training time of SVM increases exponentially, but its growth rate is lower than that of RBF. This validates its learning abilities and modeling efficiency.

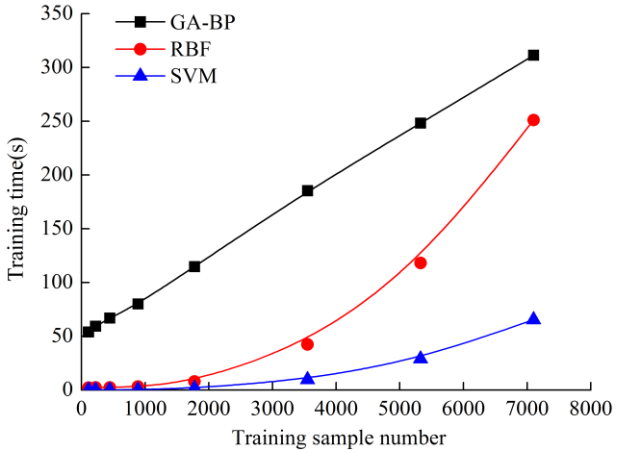

FIG.7 Training time with GA-BP, RBF and SVM forecasting models

c. Adaptability to climates

Fig.8 shows the prediction RMSE of 12 months of GA-BP, RBF and SVM in three wind farms. It is one month/season training, and one month/season prediction with error calculation. The variation trends in each figure are consistent. It shows that the prediction accuracy in spring (Mar., Apr., May) is lower than other seasons. This might be because that the weather in spring fluctuates frequently that the NWP accuracy is reduced. Meanwhile, the power output is higher due to abundant wind resources in spring, and this increases prediction error as well. Comparatively, RBF and SVM have good prediction accuracy in spring, which show their preferable generalization abilities. RBF does not have a higher demand for NWP accuracy, showing better adaptability in WF3. GA-BP gets a better predictive effect in Jun. and Aug. because of the small and stable wind conditions in summer.

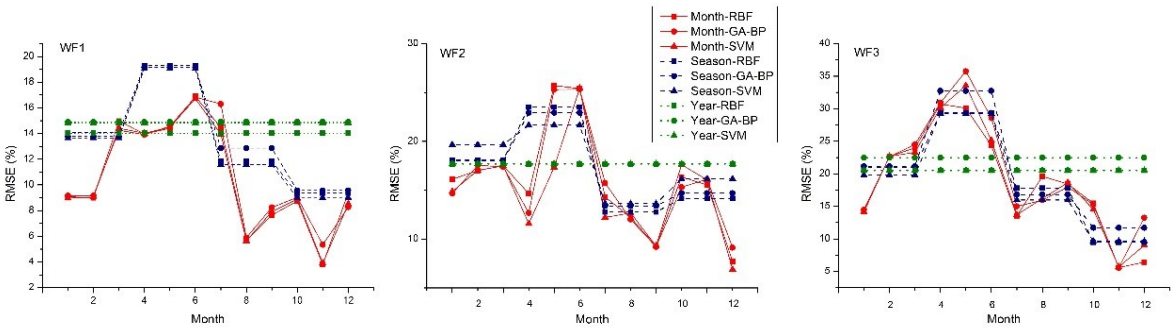

FIG. 8 Prediction RMSE of 12 months in three wind farms

d. Variability effects on predictability

Fig. 9 shows the linear dependency between three forecasting RMSEs (or predictability) and normalized wind power variability in three wind farms. The RMSE is based on forecast horizon of 24 hours. In Eq. (16), the wind power variability is defined by the mean absolute 15 minute gradient of the measured power time series P(t) that serves also as training targets[53]. It is clear to see that larger variability leads to lower predictability, which reflects the seasonal effects on forecasting error distribution.

$$Variability = \left| P\left(t + 15\,\mathrm{min}\right) - P\left(t\right) \right| \tag{16}$$

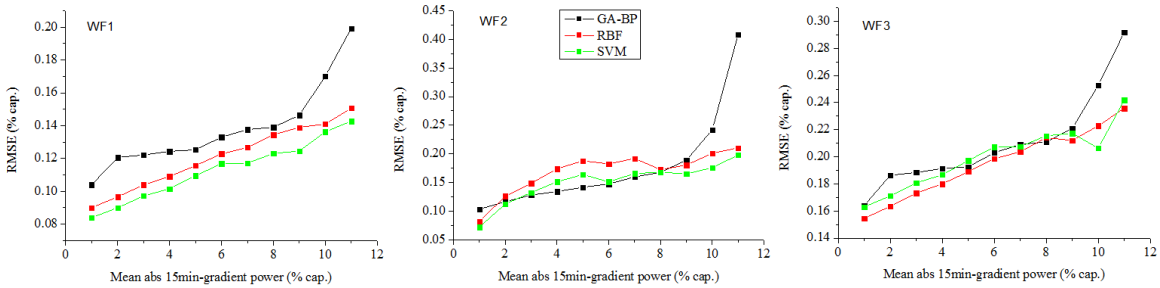

FIG. 9 Correlation between variability and predictability (RMSE) in 12 months.

e. Adaptability to the terrains

The average RMSE of every model is shown in Table II. The three models meet the requirement of RMSE less than 20%. SVM has better prediction average accuracy of 10.54% in WF1 and 14.87% in WF2. This is because that both wind farms are flat and have complete met mast data as well as better NWP data. Due to the topological complexity of WF3 and lack of partial met mast data, the prediction performance of three models is all a little worse. Relatively, RBF has better prediction accuracy of 18.76% in WF3, owing to its strong adaptability.

On the whole, RBF and SVM have a more accurate prediction than GA-BP. SVM is suitable for the wind farms with flat terrains and precise NWP data, especially coastal areas, whilst RBF shows advantages in those with complex terrains and worse NWP data.

Table II Monthly average RMSE of GA-BP, RBF and SVM models in WF1, WF2 and WF3

Model	WF1	WF2	WF3
	RMSE/%	RMSE/%	RMSE/%
GA-BP	10.92	15.87	19.99
RBF	10.67	15.98	18.76
SVM	10.54	14.87	18.90

Fig.10 shows the predictive and actual value of wind farm power output with the three models over one random day (288 time moments). Except for GA-BP model in WF2 shown in Fig.10 (b), other prediction curves are primarily in accordance with the actual value, and the error between the predictive and the actual power output is small. However, three models, especially GA-BP, cannot track the sharp power changes accurately. This is related to the features of model learning as well as the NWP accuracy. For GA-BP, it has the drawback of becoming trapped into the local minimum. Although it can be improved by GA, it requires further optimization.

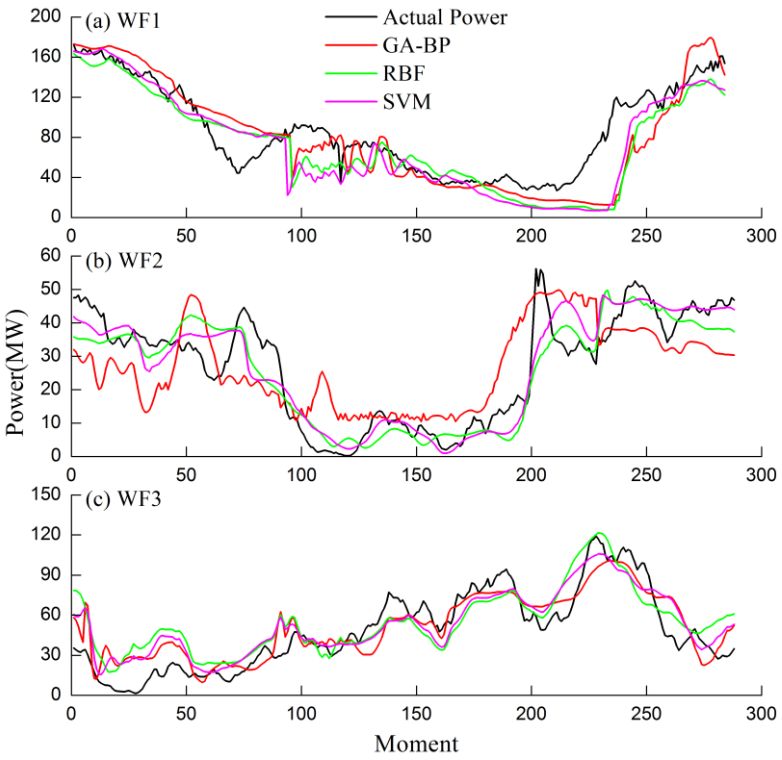

FIG.10 24-h horizon predictive and actual value of wind farms power output in wind farms on one random day in (a) WF1, (b) WF2, and (c) WF3.

VII. CONCLUSION

In this paper, GA-BP, RBF and SVM methods are established and analyzed for prediction accuracy, computation efficiency and model adaptability. Based on the case study in China, following conclusions can be drawn:

(1) Training sample effects. GA-BP and RBF have similar interaction effects from training sample number on forecasting accuracy. The prediction accuracy increases with the increase in training sample number. The change of sample number is less sensitive than that of SVM. GA-BP forecasting accuracy is equally sensitive to all size of training samples. With small training sample, the rising slope of RBF red curves is far greater than that of GA-BP. But with more samples, RBF accuracy remains constant. The prediction accuracy of SVM increases with the decreasing of training samples, however, it reduces rapidly with large training samples.

(2) The size of the training sample for different forecasting algorithm has such a distinctive influence on forecasting performance. According to the training sample analysis in this paper, it is suggested that:

- At least 7000-8000 samples should be used for GA-BP model training.
- At least 800-1000 samples should be used for RBF model training to obtain an accuracy-efficiency balanced result. Around 7000 training samples can produce better accuracy, but the effect is marginal and the training time is increased.
- At least 400-600 samples should be used for SVM model training.

- To convert from sample number to calendar time, one has to use the proper sampling time and subtract the amount of unusable data due to missing measurements, invalid data, maintenance and curtailment, and 10-min samples would work the same way as hourly samples.

(3) Model adaptability to the terrains. SVM is suitable for the wind farms with flat terrains and precise NWP data, especially coastal areas. Whilst RBF shows advantages in those wind farms with complex terrain and worse NWP data.

(4) Model adaptability to climates. RBF and SVM have good prediction accuracy in spring, while GA-BP is more suitable for summer. It is suggested to combine these three models depending on different seasons to increase the wind power prediction accuracy.

(5) Model computation efficiency. SVM modeling requires less training samples and better NWP with high computational efficiency. It is suitable for the modeling of new-built wind farms or areas with various weather conditions. The RBF model is not suitable for training samples with too much data; otherwise its computational efficiency is too low. GA-BP with lower computational efficiency demands for larger training samples making it unsuitable for the modeling of new-built wind farms or areas with a variety of weather conditions.

(6) Overall, RBF and SVM have better prediction performance than GA-BP, and prediction models which are established with one month samples at a time can improve short-term wind power prediction accuracy.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (51376062) and technology project from State Grid Corporation of China “Research on wake effect in large-scale wind farm base” and technological project from State Grid Corporation of China (国家电网科技项目). The first author is currently doing joint-educated Ph.D. program and sponsored by China Scholarship Council in University of Bath, UK.

REFERENCES

[1] GWEA: Global statistics: http://www.gwec.net/global-figures/graphs/

[2] See http:// www.nea.gov.cn/2015-02/12/c_133989991.htm for Chinese National Energy Administration:

[3] M. G. De Giorgi, A. Ficarella, and M. Tarantino, “Error analysis of short term wind power prediction models,” Appl. Energy **88**, 1298-1311(2011).

[4] G.Y.Zhang, Y.G. Liu, and Y.Q. Liu, “An advanced wind speed multi-step ahead forecasting approach with characteristic component analysis,” J. Renewable Sustainable Energy **6**, 053139(2014).

[5] X. Yang, Y. Xiao, and S. Chen, “Wind speed and generated power forecasting in wind farm,” Proc. Chin. Soc. Electr. Eng. **25**, 1-5(2005).

[6] H. Liu, E. Erdem, and J. Shi, "Comprehensive evaluation of ARMA arch (-m) approaches for modeling the mean and volatility of wind speed," Appl. Energy **88**, 3, 724–732(2011), available online at http://www.sciencedirect.com/science/article/pii/S0306261910003934.

[7] R. G. Kavasseri and K. Seetharaman, "Day-ahead wind speed forecasting using f-ARIMA models," Renew. Energy **34**, 5, 1388–1393(2009).

[8] P. Louka, G. Galanis, N. Siebert, G. Kariniotakis, P. Katsafados, I. Pytharoulis, and G. Kallos, "Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering," J. Wind Eng. Ind. Aerodyn. **96**, 12, 2348–2362(2008).

[9] E.A. Bossanyi, "Short-term wind prediction using Kalman filters," Wind Engineering. **9**, 1-8(1985).

[10] Y. A. Katsigiannis, A. G. Tsikalakis, P. S. Georgilakis, and N. Hatziargyriou, " Improved Wind Power Forecasting Using a Combined Neuro-fuzzy and Artificial Neural Network Model," Adv. Art. Intell. **3955**, 105-115(2006).

[11] S. Singh, T. S. Bhatti, and D. P. Kothari, "Wind power estimation using artificial neural network," J. of Energy Eng. **133**, 46 - 52 (2007).

[12] R. Jursa and K. Rohrig, "Short-term wind power forecasting using evolutionary algorithms for the automated specification of artificial intelligence models," Int. J. Forecasting **24**, 694-709(2008).

[13] G. Li and J. Shi, "On comparing three artificial neural networks for wind speed forecasting," Appl. Energy **87**(7), 2313–2320 (2010).

[14] Z. H. Guo, J. Wu, H. Y. Lu, and J. Z. Wang, "A case study on a hybrid wind speed forecasting model using BP neural network," Knowl. -Based Syst. **24**, 1048–56 (2011).

[15] G. N. Kariniotakis, G. S. Stavrakakis, and E. F. Nogaret, "Wind power forecasting using advanced neural network models," IEEE. Trans. Energy.Convers. **11**, 762-767 (1996).

[16] L. J. Ramirez-Rosado, L. A. Fernandez-Jimenez, and C. Monteiro, "Artificial neural network models for wind power short-term forecasting using weather predictions," in *Proceedings of the 25th IASTED International Conference on Modeling, Identification and Control*, Spain, 2006, pp. 128-132, see http://actapress.com/PaperInfo.aspx?PaperID=23039.

[17] H. W. Peng, F. R. Liu and X. F. Yang, "Short term wind power forecasting based on artificial neural network," Acta Energiae Solaris Sinica **32**, 1245-1250(2011).

[18] J. P. S. Catalão, H. M. I. Pousinho, and V. M. F. Mendes, "Short-term wind power forecasting in Portugal by neural networks and wavelet transform," Renewable Energy **36**, 1245–1251(2011).

[19] N. Amjady, F. Keynia and H. Zareipour, "Short-term wind power forecasting using ridgelet neural network," Electr. Power Syst. Res. **81**, 2099-2107 (2011).

[20] Y.Q. Liu, J.Shi, Y.P. Yang, and S. Han, "Piecewise support vector machine model for short-term wind-power prediction," Int. J. of Green Energy 6, 479-489 (2009).

[21] X. L. Wang and M. W. Wang, "Short-term wind speed forecasting of wind farm based on

least square-support vector machine," Power Syst. Technol. **34**(1), 179-184 (2010).

[22] H. Bouzgou and N. Benoudjit, "Multiple architecture system for wind speed prediction," Appl. Energy **88**, 2463–2471 (2011).

[23] Y.Q. Liu, J. Yan, S. Han, D. Infield, D. Tian, and L.Y. Gao, "An optimized short-term wind power prediction method considering NWP accuracy," Chin. Sci. Bull. **59**(11), 1167-1175 (2014).

[24] J. Yan, Liu, Y. Q. Liu, S. Han, and M. Qiu, "Wind power grouping forecasts and its uncertainty analysis using optimized relevance vector machine," Renewable and Sustainable Energy Rev. **27**, 613-621(2013).

[25] L. Li, Y. Q. Liu, Y. P. Yang, and S. Han, "Short-term wind speed forecasting based on CFD pre-calculated flow fields," Proc. Chin. Soc. Electr. Eng.. **33**, 27-32 (2013).

[26] L. Landberg, "A mathematical look at a physical power prediction model," Wind Energy **1**, 23-28 (1998).

[27] L. Landberg, "Short-term prediction of local wind conditions," J. Wind Engineering and Industrial Aerodyn. **89**(3), 235-245(2001).

[28] G. Giebel, J. Badger, I. Martí Perez, P. Louka, G. Kallos, A.M. Palomares, C. Lac, and G. Descombes, "Short-term forecasting using advanced physical modeling-the results of the ANEMOS project," in Proceedings of European Wind Energy Conference. Athens, Greece, (2006).

[29] S. Al-Deen, A. Yamaguchi and T. Ishihara, "A physical approach to wind speed prediction for wind energy forecasting," in *The Fourth International Symposium on Computational Wind Engineering*. Yokohama, Japan, (2006).

[30] S. Feng, W. Wand and C. Liu, "Study on the physical approach to wind power prediction," Proc. Chin. Soc. Electr. Eng. **30**(2), 1-6(2010) (in Chinese).

[31] S. Alessandrini, S. Sperati, and P. Pinson, "A comparison between the ECMWF and COSMO Ensemble Prediction Systems applied to short-term wind power forecasting on real data," Appl. Energy **107**, 271-280(2013).

[32] T. Hong, P. Pinson, S. Fan, "Global Energy Forecasting Competition," International Journal of Forecasting **30**(2), 357-363(2014).

[33] G. Giebel, G. Kariniotakis, and K. Brownsword, "The State of the Art in Short-term Prediction of Wind Power. 2nd version edition," Deliverable report D-1.2 of the EU project Anemos (2011), available online at http://orbit.dtu.dk/fedora/objects/orbit:83397/datastreams/file_5277161/content.

[34] C. Monterio, R. Bessa, V. Miranda, A. Botterud, J. Wang, and G.Conzelmann, "Wind Power Forecasting: State-of-the-Art," Argonne National Laboratory, a report, 2009, ANL/DIS-10-1(2009).

[35] I. J. Ramirez-Rosado, L. A. Fernandez-Jimenez, C. Monteiro, J. Sousa, and R. Bessa, "Comparison of two new short-term wind-power forecasting systems," Renewable Energy

34(7), 1848–1854(2009).

[36] M. A. Mohandes, T. O. Halawani, S. Rehman, and A. A. Hussain, "Support vector machines for wind speed prediction", Renewable Energy **29**, 939-947 (2004).

[37] See http://rda.ucar.edu/datasets/ds083.2/#!description for CISL Research Data Archive.

[38] C. Stathopoulosa, A. Kaperonia, G. Galanisa, G. Kallos, "Wind power prediction based on numerical and statistical models," J. Wind Eng. and Ind. Aerodyn. **112**, 25-38(2013).

[39] H. G. Han, Q. L. Chen, and J. F. Qiao, "Short-term wind power forecast based on the radial basis function neural network," Neural Networks **24**, 717-725(2011).

[40] H. W. Peng, F. R. Liu and X. F. Yang, "A hybrid strategy of short term wind power prediction," Renewable Energy **50**,590-595(2013).

[41] H. Zhao and S. Guo, "Wind Speed Forecasting in China: A Review," Sci. J. Energy Eng. **3**(4-1), 14-21(2015).

[42] X. Wand, H. Li, "Effective wind speed forecasting in annual prediction of output power for wind farm," Proc. Chin. Soc. Electr. Eng. **30**(8), 117-122(2010).

[43] H. Zhang, J. Zeng, "Wind speed forecasting model study based on support vector machine," Acta Energ. Sol. Sin. **31**(7), 928-932(2010).

[44] H. Yang, S. F. Gu, M. D. Cui, "Forecast of short-term wind speed in wind farms based on GA optimized LS-SVM," Power Sys. Prot. Control **39**(11), 44-48(2011).

[45] X. Y. Yang, B. J. Sun, X. F. Zhang, "Short-term wind speed forecasting based on support vector machine with similar data," Proc. Chin. Soc. Electr. Eng. **32**(4), 35-41(2012).

[46] M. Xu, X. C. Ji, J. H. Wu, M. Zhang, "A modified BP algorithm for LDPC decoding based on minimum mean square error criterion," IEICE Trans. Commun. **E93.B**(5), 1256-1259 (2010).

[47] M. D. Buhmann, *Radial basis functions: theory and implementations*, Cambridge Monographs on Applied and Computational Mathematics, (Cambridge University Press, 2009)

[48] N. Sundararajan, P. Saratchandran, Y. W. Lu, *Radial basis function neural networks with sequential learning: MRAN and its applications*, (World scientific press, Singapore, 1999).

[49] S. Mukherjee, E. Osuua, and F. Girosi, "Nonlinear prediction of chaotic time series using signal processing," in Proceedings of the IEEE Neural Networks for Signal Processing [1997] VII. Proceedings of the 1997 Workshop, 24-26 Sep 1997, Amelia Island, FL, 511-520IEEE NNSP (1997).

[50] V. Vapnik. *Statistical learning theory* (Wiley, New York, 1998).

[51] P. Pinson, H. Nielsen, H. Madsen, and G. Kariniotakis, "Skill forecasting from ensemble predictions of wind power," Appl. Energy **86**, 78, 1326–1334(2009).

[52] H. Madsen, P. Pinson, G. Kariniotakis, H. A. Nielsen, and T. Nielse, "Standardizing the performance evaluation of short term wind power prediction models," Wind Eng. **29**(6), 475-489(2005).

[53] Jan Dobschinski, “How good is my forecast? Comparability of wind power forecast errors,” in *13th International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plants*, Berlin, Germany, 2014, pp. 1-6.